Python for Finance and Algorithmic Trading

Machine Learning, Deep Learning, Time Series Analysis, Risk and Portfolio Management for Metatrader™ 5 Live Trading

POWERED BY

jupyter

Quantreo

Lucas Inglese

Book Link : https://www.amazon.com/gp/product/B09HG18CYL

EXTRACT Chapter 7: ARMA

7.1.3. The linear regression

In this subsection, we are not going to create a linear regression model. Indeed, we will explain how linear regression works because this concept is necessary to understand the ARMAs models.

Linear regression is a predictive algorithm. Indeed, the objective of this algorithm is to allow us to predict some data. For example, if you know a relationship between gold and S&P 500, you can use linear regression to predict the S&P 500 stock price using the gold price.

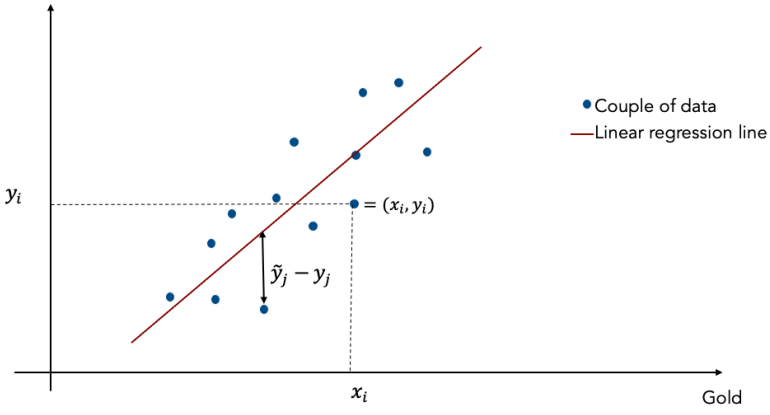
Let us see how to compute a linear regression using an optimization problem and explain more profoundly the intuition using figure 7.5.

$$min_{\beta} \sum (\tilde{y}_i - y_i)^2$$

Where $\tilde{y}_i$ is the predicted value, y_i the real value and β are the parameters of the model.

The intuition of the linear regression can be resumed by an algorithm that tries to find the best way to minimize the distance between the predicted and real values.

Figure 7.5: Linear regression intuition



This figure shows how linear regression works with a model that takes the gold and predicts the S&P 500 price.

It is important to understand that for a dataset X which has m rows and n columns $\tilde{y}_i = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_n x_{i,n}$.In our example with S&P 500 and gold $\tilde{y}_i = \beta_0 + \beta_1 \, gold \, price_i$ (β_0 to β_n are the model's parameters).

EXTRACT Chapter 8: Linear Regression

8.1.1. Reminder about regression

In the previous chapter, we have done a little resume of the linear regression. Here, we will do a little reminder of this and explain how regression works for any regression algorithms.

- All regression models want to predict a continuous value. For example, if a model predicts the percentage of stock variation, the prediction will be contained into the interval [-30%; +30%],

and it can take all values into this interval (for example, 1.57%, 6.78%, 13.39%).

- Intuitively, the idea of regression is to find a model which can be closer to each value of a dataset. We can compute the sum of the distance between the predicted value $\tilde{y}_i$ and the real value y_i using the mean squared error (MSE). Mathematically, MSE = $\sum(\tilde{y}_i - y_i)^2$

- All regression algorithms have their method to be optimized. However, all can be test with the MSE to find if it is a good model or not. In comparison, the line regression is the only which optimizes the MSE to find the best parameters.